

***4. International Congress of Dialectologists and Geolinguistics, Riga, 2003***

**L'interprétation dans les atlas linguistiques? Le cas de l'Atlas linguistique  
basque (EHHA). Application à la carte 'caille'.**

Gotzon Aurrekoetxea  
Université du Pays Basque,  
E-01006 VITORIA-GASTEIZ

Xarles Videgain  
Université de Pau et des Pays de l'Adour  
Faculté pluridisciplinaire  
F-64100 Bayonne

0. Résumé

L'atlas linguistique basque, en cours de publication, qui suit la plupart des présupposés de la géolinguistique traditionnelle, apporte cependant quelques nouveautés méthodologiques, en particulier dans la présentation des données : responsables, index, cartes à petit format des formes lemmatisées et non des réponses recueillies, sur support papier et sur CD-rom. Ces nouveautés sont justifiées par le public recherché, que nous souhaitons plus vaste que celui des seuls spécialistes. Elles permettent aussi une meilleure mise en évidence de la fragmentation dialectale en montrant divers cas de figure sur les cartes dont la sémiologie sera exposée à travers divers exemples. Une illustration à partir de la carte 'caille' / 'quail' montrera que cette cartographie n'épuise pas les informations conservées dans la base de données. Cette contribution associe donc une vue d'ensemble sur la présentation des données d'atlas à une étude plus particulière sur des données précises.

1. L'Atlas Linguistique Basque (EHHA) parmi les atlas linguistiques

L'Atlas Linguistique Basque (EHHA) est un projet tardif dans l'atlantographie européenne.

Les raisons sont multiples : la répression contre la langue basque, l'inexistence d'une université du Pays Basque jusqu'à pratiquement 1980, la précarité des études sur la langue basque, la pénurie de chercheurs professionnels sur le domaine basque.

Mais les tentatives en faveur d'un atlas du basque datent déjà des années 20. Il serait trop long de retracer la liste des projets de réalisation de l'atlas, mais ce fut l'Académie de la langue basque-Euskaltzaindia qui commença les travaux de l'atlas basque avec la création d'une commission spéciale en 1983 et l'obtention des moyens nécessaires à sa réalisation.

L'EHHA est donc une entreprise tardive, si on la considère en relation avec les autres projets atlantographiques d'Europe ; il est en retard aussi pour la collecte convenable des variétés de la langue basque et de la culture traditionnelle, que d'autres atlas ont menée à bien il y a longtemps. Mais il n'est pas encore trop tard, pour la collecte des témoignages des variétés idiomatiques menacées à court terme et celles relatives aux travaux propres à la culture traditionnelle (labour, pêche etc). C'est pour cela que l'EHHA a accordé une attention toute particulière à recueillir les vestiges et souvenirs des traditions anciennes. Ce choix délibéré en faveur de la situation la plus ancienne du basque parmi les locuteurs actuels n'en contraste pas moins avec l'utilisation des technologies plus avancées pour le traitement des données et leur publication.

La première présentation du projet à la communauté scientifique internationale se fit dans le cadre d'un *Congrès International de Dialectologie* (1991), tenu à Bilbao en l'honneur de Louis Lucien Bonaparte, considéré comme le père de la dialectologie basque, et ce à l'occasion du centenaire de son décès (pour plus de détails voir Aurrekoetxea, 1992).

La présentation des premières données et la version beta du CD ont été exposées à Bayonne dans le cadre du workshop « Corpus et traitement informatisé en géolinguistique » (22-23 Juin 2001), organisé par la European Science Foundation. Ce fut alors une présentation des premières données du projet entre chercheurs qui travaillent sur l'atlantographie, et dont quelques uns sont ici présents.

Aujourd'hui, nous voulons présenter, dans le cadre du « 4. Congress International of Dialectologists and Geolinguistics », l'argumentation qui nous a fait choisir un certain type de carte et préciser pourquoi nous proposons ce type de carte et non pas les cartes classiques des atlas linguistiques traditionnels.

Cette réflexion rejoint le souci de tous les auteurs d'atlas de publier les résultats d'enquête de la manière la plus scientifique, mais nos conclusions dans ce domaine nous ont amené à nous éloigner, sur certains points, de certains de nos prédécesseurs.

À dire la vérité, au début, quand nous étions en train de décider le modèle des cartes linguistiques à adopter, nous craignions quelque peu de nous éloigner de la publication typique, mais le format des volumes traditionnels ainsi que le travail cartographique insuffisant à notre avis dans les atlas traditionnels, difficulté d'utilisation des données, nous poussaient à retravailler la question de la publication des données pour qu'elles soient de consultation plus facile, que leur consultation puisse se faire à plusieurs niveaux et soit accessible au public cultivé plus vaste que les seuls spécialistes.

Nous doutions donc de l'utilité des publications d'atlas-cartes à la stricte manière gilliéronienne. Nous nous demandions pourquoi il n'était pas possible de publier dans une unique publication aussi bien les données « brutes » qu'une tentative d'ordonnancement de ces données, comme l'a fait par exemple Bernard Tanguy (1984), avec les données des *Atlas linguistiques de la Basse-Bretagne* de Pierre Le Roux et *l'Atlas linguistique de la France* de J. Gilliéron.

Deux refus de notre part nous ont aussi poussé dans nos choix :

- Non aux cartes des réponses brutes (avec copie des réalisations phonétiques sur le point concerné)
- Non aux volumes de très grand format.

L'approfondissement des points déjà exposés et ces refus initiaux nous donnaient deux lignes à suivre : a) d'une part, une élaboration linguistique des données s'imposait ; b) d'autre part, une élaboration cartographique nouvelle était à choisir. Nous soulignons que plus qu'une interprétation, il s'agit d'une élaboration ou d'un ordonnancement des données.

### 1.1. Élaboration linguistique et cartographique

Il faut dire en effet que cette élaboration linguistique est en fait une classification typologique des données (nous parlerons ici du lexique). S'il y a interprétation c'est uniquement en cela, par un effort typologique, tel qu'une absence

ou présence d'un morphème, absence ou présence d'une caractéristique phonétique (palatalisation, amuïssement, sonorisation, etc.), une variation lexématique, etc.

Au moyen de cette élaboration nous sommes en état de présenter les données recueillies (à partir des phonogrammes en enquête) selon un alphabet phonétique et en alphabet orthographique selon une translittération automatisée construite par nos soins. Les données en alphabet IPA seront publiées dans l'atlas sur des listes ordonnées par points d'enquête. C'est le niveau le plus précis de lecture des données. Nous pensons qu'il n'y a pas réellement besoin de cartes pour publier les données phonétiques. La carte n'est pas non plus nécessaire pour opérer une recherche ou requête phonologique ou phonétique dans le corpus des données collectées.

#### a) Données brutes

Nous entendons qu'il est nécessaire de donner au lecteur les données recueillies brutes, avec la plus grande fidélité possible et avec l'unique interprétation du transcripateur. Comme dit ci-dessus ces données seront publiées dans une liste des points d'enquête ordonnés alphabétiquement (pour des raisons financières et de temps, nous n'avons pu donner un atlas parlant c'est à dire que nous ne fournissons pas encore le son correspondant à la forme recueillie, comme le fait remarquablement l'Atlas ladin par exemple).

#### b) Données élaborées : lemmatisation

Les données placées en alphabet phonétique dans la base des données, ont été transférées d'abord, sous forme orthographique, de façon automatique. Pour chaque réponse qu'on informatise phonétiquement le système nous donne sa version orthographique. En partant de ces données orthographiques on réalise la lemmatisation (même si théoriquement, et la machine le permet, on peut aussi lemmatiser les réalisations phonétiques). Cette lemmatisation est une classification des données selon divers critères. De ce point de vue il ne faut pas entendre la lemmatisation au sens aussi strict qu'en lexicologie pure. D'autres critères que ceux purement lexicaux sont en effet pris en ligne de compte.

Ces critères sont donc le plus souvent: a) étymologie ; b) absence ou présence d'un morphème ; c) absence ou présence d'un phonème ou son déterminé ; d) la

présence ou absence d'une règle phonologique ; e) emprunt, de datation variable, d'intégration plus ou moins assumée.

Cette lemmatisation on fait avec un lemmatisateur construit « ad hoc » par l'équipe d'EHHA. Le dialectologue part d'une liste des réponses orthographiques (en haut à gauche) qu'il réduit, selon les critères déjà signalés, à un maximum de 16 superlemmes (en bas à droite).

Cette typologisation des données peut obéir quelque fois à des critères plus ou moins subjectifs surtout dans le poids relatif que l'on donne à l'un ou l'autre de ces critères mais elle est toujours une classification des données sur la variation linguistique. C'est en cela que nous parlons plutôt d'« élaboration », et non d'« interprétation ».

Nous que dans l'interprétation des données on se pose le pourquoi de tel phénomène linguistique en fonction de l'histoire, de la géographie, de l'organisation politico-administrative, de la culture, etc.

Dans notre élaboration linguistique nous faisons abstraction de ces disciplines ou points de vue. Elles n'interfèrent pas directement dans la lemmatisation et pourront être objet de recherches ultérieures. Nous ne nous demandons pas la raison de l'apparition d'un mot dans un endroit déterminé. Nous procédons simplement à la constatation d'un fait linguistique, et voulons qu'il soit à la fois visible et lisible.

En effet, cette lemmatisation est absolument nécessaire pour donner au public des données plus lisibles et plus faciles à assimiler : elle n'est donc pas une véritable interprétation des données.

A l'autre bout, même pour un lecteur spécialiste de la langue, la publication point par point sur la carte de toutes les variantes recueillies phonétiquement ne nous semble pas la solution préférable ; nous craignons que dans ce cas les arbres ne nous laissent pas voir la forêt.

La lemmatisation se fait question par question et peut atteindre des critères très différents. Dans une question avec une grande variété de lexèmes on traitera surtout la variation lexicale et suffixale ('bosque' / 'forêt') ; mais dans une carte monolexématique ('rama' / 'branche') on peut davantage mettre en valeur un aspect phonologique, dans ce cas la réalisation de l'occlusive intervocalique (*adar*, *agar*, *abar*, *aar*) ; ou dans la carte 'sequía' / 'sécheresse' l'accommodation de l'emprunt: *sikete*, *siketa*, *sikute*, *sekate*, *segate*, *sekura*) ou l'apparition du son [ü] a l'Est du territoire, dans la carte « pulga / puce ».

Cette lemmatisation se veut donc une aide pour voir plus facilement et plus aisément les données recueillies. Et accessoirement elle permet de rendre les données sur un format A4 plus raisonnable que les cartes de grandes dimensions des atlas traditionnels.

Parfois, il nous apparaît que deux degrés de lemmatisation sont nécessaires sur une même question : nous pouvons recourir à deux cartes, celle habituelle et une carte de plus petite dimension sur laquelle nous ne pouvons mettre en valeur que quelques traits discriminants et que se trouvent dans le coin droit bas de la page des réponses.

Ici non plus il n'y a pas d'interprétation à proprement parler cartographique dans les cartes de l'atlas basque. La cartographie de beaucoup d'atlas linguistiques est à notre avis d'une sémiologie très déficiente. Ces atlas nous apparaissent souvent comme très peu lisibles et indéchiffrables. Il faut un temps bien trop long pour organiser, comprendre, analyser les données recueillies sur un quelconque objet linguistique. Pourtant le travail de collecte y a été remarquable.

Sans perte de qualité, nos publications, selon nous, doivent être plus attractives et convaincre un public plus ample que celui des spécialistes.

Les raisons financières plaident pour l'abandon des grands formats et pour l'élaboration de dimensions plus réduites. Le grand format est cher, peu pratique.

### c) Les cartes

Le fonds de carte d'EHHA recourt aux triangles de Thiessen, ou dits aussi de Voronoï, déjà largement utilisés en dialectométrie par H. Goebl et d'autres utilisateurs (voir par exemple Goebl, 1992). En conséquence, les zones d'implantation sur la carte ne sont pas un point mais un polygone. Leurs contours ne sont pas les frontières géographiques de chaque commune enquêtée mais une modélisation. Ainsi, il n'existe pas de trous blancs entre les communes enquêtées. L'avantage, à notre avis, est de faire se dégager des aires linguistiques. En effet, une couleur est assignée à chaque item de la lemmatisation. Les couleurs choisies sont dénuées de toute valeur symbolique et sont indépendantes entre elles d'une carte à l'autre. L'absence de réponse est traduite par la couleur blanche.

C'est sur espace et ce format que nous avons pensé un système cartographique que nous permet l'introduction d'un maximum de trois réponses dans chaque commune par question. En effet, la méthodologie d'enquête suivie dans l'EHHA nous

a fourni des réponses qu'on peut classer en deux groupes : les réponses spontanées d'une part et les réponses proposées par l'enquêteur et assumées par les témoins comme propres de l'endroit enquêté (Ravier, 1965).

Dans la cartographie que nous proposons nous avons la possibilité de représenter par chaque point d'enquête deux réponses spontanées sur la carte : la première en couleur soutenue et la deuxième en couleur tramée. En outre, les réponses obtenues sur sollicitation ou réponses proposées, réponses de rang 2, sont représentées par un cercle.

A notre avis, nous avons donc obtenu la représentation des données sur un espace réduit et en allant jusqu'à trois réponses par point. Nous aidons à la lisibilité des espaces distingués par la couleur en écrivant sur l'espace concerné le lemme correspondant.

Malgré l'exigüité du format, nous avons la possibilité de faire une deuxième carte plus petite pour chaque question. Nous l'avons déjà dit, cette deuxième carte, placée près du responsable, ne met en valeur que quelques traits choisis par l'auteur de la carte.

Notre intention est donc que cet atlas soit utile et utilisable, plus que bien des atlas dont la visite est un parcours d'obstacles qui peut rebuter.

## 1.2. La publication sur papier

### Les index

Un de nos objectifs déclarés est l'utilisation rapide et aisée de l'atlas. A la question triviale « tel mot a-t-il été recueilli dans l'atlas ? » nous voulons pouvoir répondre immédiatement. A l'instar d'autres atlas (mais à la différence de la plupart des atlas anciens) nous donnons divers index des items recueillis, sous la forme orthographique. Nous donnons aussi les arbres de lemmatisation, c'est-à-dire la façon dont la carte a été élaborée quant à la classification de toutes les réponses en lemmes.

## 1.3. La publication en CD. Donner tous les données aux chercheurs :

La publication en CD donne deux possibilités. D'abord la possibilité de la consultation de toute l'information qu'on publiera dans le livre : réponses, cartes, index. Mais elle ouvre une autre possibilité. D'une part le lecteur pourra réaliser sa

propre carte à la lumière des données recueillies. Il peut aussi lire les minutes d'enquêtes avec les textes des locuteurs et des enquêteurs.

## 2. Illustration.

La carte 'caille'.

Si nous avons présenté jusqu'ici les principales caractéristiques de la cartographie de l'atlas basque, nous voudrions maintenant illustrer ce propos à partir d'une seule question qui porte le numéro 05430 dans notre questionnaire d'atlas et qui porte sur l'oiseau 'caille' / 'codorniz' (*Coturnix coturnix*) et ce pour montrer que la cartographie adoptée ne saurait épuiser l'ensemble des informations recueillies dans l'atlas.

Nous allons d'abord rappeler ce que la cartographie de l'atlas révèle aisément au lecteur. En effet, la carte 'caille' de l'atlas linguistique basque permet plusieurs degrés de lecture et de compréhension de la fragmentation dialectale.

D'une part, nous donnons une liste complète des dénominations recueillies pour la 'caille'. Chaque réponse est donnée dans sa réalisation phonétique en alphabet IPA dans un responsive sous forme d'un tableau prenant toute la page (et non pas sur une carte point par point sur laquelle seraient reportées les formes phonétiques sur l'espace correspondant au lieu d'enquête).

D'autre part, l'atlas donne la carte en couleurs qui, après lemmatisation, situe les réponses relevées: dans ce cas précis, la carte 'caille' révèle à gros traits quatre espaces portant quatre couleurs distinctes, les quatre aires étant de taille très diverse.

Sur le premier espace, on constate un emprunt généralisé en domaine oriental avec *kalla* (du français 'caille' et gascon 'calle'), l'isoglosse suivant la frontière politique et caractérisant en effet un espace en contact avec le français et le gascon.

Sur le deuxième espace, on trouve une forme motivée *galeper* composé de *gari+eper* littéralement 'perdrix des blés', avec quelques variations sur le timbre vocalique: *galeper*, *galaper*, *galiper*, *galapar*.

Sur un troisième espace on trouve une autre forme, que nous expliquerons plus loin, le mot *poxpolin*.

Enfin, dans le quatrième espace, une zone réduite est présente autour du vocable *ttorttola*. Si la zone ne comprenait pas plusieurs communes enquêtées, on pourrait penser à une erreur des enquêteurs ou des témoins puisque *ttorttola* est un



emprunt du castillan ‘tórtola’ qui désigne non pas la ‘caille’ mais la ‘tourterelle’. Il faut se reporter à la carte ‘tourterelle’ pour voir si ce même terme de *ttortola* revient aussi, auquel cas il faudrait considérer ici les réponses comme très douteuses, car il est assez peu probable que le même terme désigne durablement deux oiseaux aussi distincts que caille et tourterelle.

La partition de la carte est assez simple dans le cas qui nous occupe (d’autres cartes poussent à des choix plus délicats en termes de lemmatisation) mais à qui voudrait connaître le détail des choix faits par l’éditeur de la carte il suffirait de se reporter en fin d’ouvrage à l’arbre de lemmatisation qui permet de suivre l’itinéraire du réalisateur de la carte, puisque chaque superlemme comprend l’ensemble des réponses qui ont été regroupées sous la même couleur.

Le lecteur attentif tente de comprendre la fragmentation dialectale en premier lieu à partir de la carte. Il recourt alors à son savoir philologique mais il peut aussi accéder dans le cas de notre atlas aux informations données par les témoins d’enquête. En effet la base de données fournit les minutes d’enquête: toutes les enquêtes ont été enregistrées au magnétophone sur bandes magnétiques, lesquelles bandes ont été transférées sur CD, soit un trésor de plusieurs milliers d’heures d’enregistrement. Les minutes d’enquête ne donnent pas le suivi de tout le détail de l’enquête mais permettent cependant de relever bon nombre de phénomènes.

Dans le cas présent, le savoir philologique suffit ici pour l’examen de trois aires de la ‘caille’: celles de *kalla*, *galeper*, *ttortola*. Certes on admettra l’affirmation suivante: “En principe en étymologie traditionnelle, le secours de l’intuition native n’est pas sollicité” (Chambon 1989). Mais rien n’interdit de consulter dans la base ou dans les phonogrammes les commentaires des locuteurs (ces commentaires ayant été plus ou moins fidèlement retranscrits ou résumés par enquêteur et transcripteur). Cette possibilité devient plus nécessaire encore si l’on adopte un point de vue ethnolinguistique: “expliquer les faits linguistiques par le biais de documents folkloriques” (Dinguirard, 1976).

Ce point de vue peut nous aider pour examiner la zone où *poxpolin* désigne la caille. A dire vrai, nous pensions d’abord y trouver avec Chambon un “démimologisme”, dénomination tirée d’une formule mimologique appartenant à la littérature orale. Les membres de cette classe, qu’il convient de distinguer aussi bien des “homonymes” de Buyssens que des onomatopées, relèvent d’un cas particulier de délocution nominale sur une base énonciative constituée par un fragment de

discours rapporté au style direct” (Chambon 1989). Rappelons que par “mimologismes” on entend “les formules par lesquelles on interprète plaisamment les chants des oiseaux et les cris des divers animaux” (Arnaudin ap. Chambon 1989) et le monde occitan ne manque pas de ces collectes de mimologismes (Perbosc, Rougé, Gourvest, Chadeuil...) en dehors des travaux d’Arnaudin. Ces mimologismes ne sont pas inconnus en domaine basque mais notre atlas ne les a pas recherchés de manière monographique et nous avons justement recueilli à Pagolle la formule [petedet] dans lequel on reconnaîtra la formule “Paie tes dettes”. Mais nous y avons plutôt vu un cas de diglossie, ce mimologisme n’ayant de sens qu’en français alors que la locutrice qui nous donnait cette information savait certes le français mais répondait en basque à nos questions. Nous y voyions là une transmission d’un savoir de type encyclopédique ou ethnographique intraduisible mot à mot, ce qui aurait facilité son passage en l’état du français au basque.

Ce lien entre emprunt et délocutif (forme plus générale que le démimologisme) nous était apparu aussi dans la désignation de certains objets lors de notre collecte d’atlas. C’est ainsi qu’à Amorebieta en Biscaye, les mots *benaki* et *benaka* désignent un type de filet de pêche en rivière. Bien entendu il faut y lire la séquence en castillan “Ven aquí!” ‘Viens ici!’ et “Ven acá!”, ‘Viens par ici’. La forme du verbe à l’impératif assure qu’il s’agit bien d’une injonction, nécessairement en style direct. De même en français régional à Bayonne on connaît le même type de filet sous le nom de *rape-tout*. S. Palay donne dans ce sens en gascon *rapatout* qu’il glose par une traduction mot à mot et une définition: ‘qui prend tout, un engin de pêche’. La source n’est pas un simple substantif ‘un rape-tout’ (du type ‘fait-tout’) mais sans doute une phrase ‘rape tout!’, sur le gascon *arrapà* signifiant ‘saisir, attraper, s’emparer’. Et si Palay donne une acception métaphorique de *rapatout* au sens de ‘avide, qui ramasse-tout’ en parlant de quelqu’un, cette acception n’est pas ignorée du basque qui connaît le délocutif *orobil* (littéralement ‘ramasse-tout’) mais construit sur le basque à la différence des appellations concrètes ci-dessus qui sont des emprunts.

Dans le droit fil de ces remarques, il était donc tentant de voir si le terme *poxpolin* pour la caille n’était pas un démimologisme. La lecture des commentaires des locuteurs n’a pas été vaine pour les raisons exposées ci-dessous mais nous devons avouer avoir fait buisson creux quant au démimologisme. Nulle part, en effet, le mot de *poxpolin* ne semble pouvoir être interprété comme une séquence ayant un sens.

Cependant le jeu des commentaires des locuteurs croisé avec la méthodologie des ‘faits négatifs’ nous a permis de faire contre mauvaise fortune bon cœur puisque d’autres phénomènes sont apparus qui précisent ce que la cartographie ne saurait indiquer. Nous nous intéresserons ici d’abord au mot *poxpolin* ayant pour statut d’être ou non une réponse à la question ‘caille’ et sur lequel la base de données fournit les commentaires des locuteurs. Puis nous parlerons de l’aire de *kalla*.

Nous distinguerons trois espaces distincts pour *poxpolin*. Dans le premier, nous examinerons les communes où *galeper* a été la première réponse en domaine occidental en Biscaye, *poxpolin* étant relégué à un stade inférieur. Dans le deuxième espace nous traiterons des communes dans lesquelles *poxpolin* est donné en première réponse. Dans le troisième espace, *poxpolin* est connu mais n’est pas une réponse à ‘caille’.

Examinons dans le détail la réalité des données. Dans l’espace que nous appellerons A1, nous traiterons donc des lieux d’enquête où *galeper* est la première réponse fournie: il s’agit, en domaine occidental et dialecte biscayen, des points que nous examinons maintenant.

A Getxo (point 101), la question a été posée en castillan avec le signifiant ‘codorniz’. La réponse *galeper* a été obtenue et le terme *poxpolin* a été accepté après que l’enquêteur l’ait proposé au témoin (selon la méthode des faits négatifs exposée par Van Gennep et explicitée en dialectologie par Ravier). A Zollo (point 103) en réponse au fait négatif ou terme proposé *pospolin* le témoin accepte le terme et le commente ainsi: *Len ibiltan san pospoline gari ondoetatik. / Autrefois la caille (pospolin) se déplaçait dans le blé (gari)*. Le locuteur n’établit pas de lien entre le nom de l’oiseau et son chant et il a donné déjà *galeper* comme première réponse et spontanée. A Leioa (point 102) on trouve aussi le couple réponse *galapar* (avec une légère variation phonétique *galapar* dûe à l’harmonisation vocalique) et la réponse sollicitée *pospolin*. Les autres communes comme Bakio, Zollo, Uribe-Etxebarria produisent aussi le même couple. A Uribe-Etxebarri (point 107) *galeper* a été donné spontanément, le terme proposé *pospolin* est accepté par le témoin qui signale cependant que le premier vocable est davantage utilisé. Nous ferons une remarque sur ce dernier village dans lequel le locuteur a commenté qu’il existe aujourd’hui moins de *poxpolin* (‘caille’) que de *epher* (perdrix) mais que *galeper* est plus fréquent que *poxpolin*: dans la première opposition il s’agit des référents, des oiseaux, dans une réflexion de type écologique, ou cynégétique, dans la seconde il s’agit d’une réflexion

métalinguistique. Parfois, il peut arriver, le temps lui étant compté, que l'enquêteur n'ait pas toujours clairement indiqué la nuance et il ne reste plus qu'à aller consulter le phonogramme: ce n'est pas le cas ici puisque la distinction a été sauvegardée.

Plus à l'est, Gizaburuaga (point 133) donne pour seule réponse *galipar* et n'accepte pas le terme proposé *poxpolin*. A Sondika (point 104) *galeper* est réponse donnée en premier rang puis le témoin fourni aussi *pospolin* qu'il commente: *Pospolin esaten dotzie kantetan dauelako "pop-pollin! Pop-pollin! / On l'appelle poxpolin parce qu'elle chante pop-pollin! Pop-pollin*, et l'on note la légère différence entre l'onomatopée *pop-pollin* et la désignation *poxpolin*. A Larrabetzu (point 114) *galeper* est donné comme réponse spontanée mais sur la suggestion de l'enquêteur de savoir s'il n'existe pas d'autre terme pour désigner cet oiseau, le locuteur donne de lui-même comme autre terme possible *pospolin*, où l'on voit le statut délicat du fait négatif, souvent condamné par l'impatience de l'enquêteur à n'être que fait proposé alors qu'il est peut-être une réponse, seulement moins immédiatement disponible dans un contexte donné. Enfin à Fika (point 113) et Bakio (112), les deux vocables *galeper* et *pospolin* sont fournis tous les deux comme réponses spontanées.

Si nous récapitulons les données dans cette zone biscayenne où *galeper* est cartographié en réponse de rang nous avons aperçu plusieurs degrés de relation que la lecture de la base précise:

- soit les deux termes *galeper* et *poxpolin* sont donnés comme d'usage égal, par exemple à Fika. La rayure sur la carte fait écho à ce phénomène.
- soit les deux termes sont donnés à égalité mais le témoin considère que l'un des termes est d'usage moins courant, par exemple à Uribe-Etxebarria mais la carte fournit la même sémiologie que pour le cas précédent, elle ne dit pas que l'un des termes est moins pratiqué et il vient en principe en deuxième rang.
- soit un terme est fourni immédiatement par le locuteur, puis, après un moment de réflexion, un autre terme est indiqué, ce qui suppose que l'enquêteur lui a laissé le temps de réfléchir et apporte de l'eau au moulin des enquêteurs qui préfèrent ne pas suivre à la lettre la pratique du 'premier jet' de Gilliéron. Ici encore la sémiologie de la carte est la même que ci-dessus et ne rend pas compte d'une disponibilité peut-être moins forte pour le terme produit en deuxième lieu.

La lecture de la base permet aussi de révéler l'intérêt et les limites de la méthodologie des faits négatifs, dont la puissance même peut induire en erreur. En effet, par définition, un terme proposé n'est accepté que s'il a été proposé: sous cette

lapalissade nous voulons dire que le savoir ou les présupposés de l'enquêteur ainsi que l'aspect systématique de la quête ne sont pas sans conséquence dans cette méthode. Il peut en effet s'opérer une dissymétrie qui fait que certains termes 'majoritaires' ou considérés comme plus 'standard' soient plus souvent présentés au locuteur de zones périphériques ou 'minoritaires' que l'inverse. Nous laissons cette question pour une autre étude monographique sur l'utilisation faite par les enquêteurs du stock de faits négatifs en fonction du lieu d'enquête.

Nous passons maintenant à l'espace que nous appellerons A2, toujours en domaine biscayen, où *pospolin* est première réponse sur une vingtaine de communes et seule réponse sur une quinzaine de communes. Parmi les commentaires qu'offre la base nous relèverons par exemple à Zeberio (point 111) et Elorrio (point 131) que *pospolin* est donné comme réponse spontanée et *galeper* sur proposition. Mais le témoin d'Elorrio précise *Bietara esaten dotze / On les appelle des deux manières*. En 119 à Errigoiti le témoin glose la réponse proposée *galeper* par *Gariarteko eperra, oin gariripe estéku te, onen antzérakotxue, poxpolinen antzerako txoritxue a be / La perdrix d'à travers les blés, maintenant nous n'avons plus de blé, mais elle ressemble à poxpolin, c'est un petit oiseau de même allure*. Le locuteur laisse entendre que ce n'est pas tout à fait le même oiseau mais surtout il "démonte" le terme *galeper* en *gari (arteko) eperra (la perdrix d'à travers le blé)* en mettant l'accent sur la motivation. Busturia (point 121) commente ainsi la réponse *poxpolin*: *Emen otensien len asko; or itxaso egalien egote san garixe, oooortxé. / Autrefois il y en avait beaucoup ici, en bord de mer il y avait du blé et c'est là qu'elles demeuraient*. On y voit un écart entre *gari*, blé présenté comme habitat de l'oiseau *poxpolin*. Même chose à 125 Ibarruri: *Gari artean abixe pe! Artu arrautzak eta goien atara txitak... / Même le nid était fait dans les blés, on prenait les oeufs et on obtenait les poussins*, qui marque un écart entre *gari* et *poxpolin*.

Si nous résumons les données recueillies dans cette zone, nous remarquons qu'il y a été recueilli peu de commentaires sur le mot *poxpolin* qui est la réponse et que c'est le terme proposé *galeper*, connu comme étant utilisé ailleurs que sur le point d'enquête qui est tenu comme motivé sur la construction 'perdrix des blés'. Le blé n'est invoqué par rapport à l'oiseau *poxpolin* que comme son habitat. Curieusement, très peu parlent dans la zone de *poxpolin* du chant prêté à l'oiseau.

Il est temps de passer à l'espace central sur la carte, que nous appellerons A3, qui couvre une partie du biscayen guipuscoan, le guipuscoan et une grande partie du

navarraï. Ici c'est le mot *galeper* qui est reporté en réponse de premier rang et presque partout comme seule réponse. Nous en sélectionnons quelques commentaires tirés de la base.

A Arrasate (point 302) *galeper* est la réponse spontanée avec éclaircissement de la motivation "*Onik gariten da ibliten di(r). / Ces oiseaux (galeper) vivent dans le blé (gari). Poxpolin* est accepté comme réponse proposée. Même phénomène à 305 Oñati. A Oñati Araotz (point) 304, *poxpolin* est accepté aussi comme réponse proposée mais le témoin considère que ce terme est moins utilisé que *galeper*. Seul Zestoa dans cette aire a proposé les deux réponses spontanées comme *et galeper* et *poxpolin*: on voit donc une dissymétrie avec ce qui se passe dans l'espace A1 signalé plus haut où le même cas de figure était relativement plus fréquent.

Le point 306 Bergara est important puisque le témoin déclare quand *pospolin* lui est proposé: *Pospolin esaten du kantuen entzundakuan... da kazari batek ilbaditu a! koño! galeperrak il dittu. / On dit pospolin quand on l'entend chanter mais si un chasseur les tue, ah putain il a tué des galeper. Poxpolin* nous a paru ici de statut incertain et nous l'avons laissé en réponse proposée, ce qui n'est sans doute pas le bon choix et d'autre part il faudrait recourir au son tiré du phonogramme pour savoir si l'intonation, le recours à la musique ou à une rythmique marquée, tous processus que la transcription phonétique ordinaire écrase, sont ceux d'une onomatopée ou ceux d'une désignation.

A Mendaro (point 308) c'est le témoin lui-même qui imite par la séquence *poxpoliñ poxpoliñ!* le chant de l'oiseau dont la désignation est *galeper*. Même réaction à Oresa (point 332): *Poxpoli! poxpoli! kantatzen du orrek... bi ots kra! kra! ta atzea poxpoli poxpoli... Cet oiseau chante poxpoli, deux sons kra! Kra! et de nouveau poxpoli poxpoli!*. On le comprendra, le terme *poxpolin* au sens d'onomatopée du chant de la caille ne peut pas aller sur la carte et seule la base peut nous donner une information sur la proximité sémantique de ce terme avec ce dont la carte veut rendre compte. Notre cartographie a pour règle en effet de ne pas représenter un terme proposé s'il porte les caractéristiques suivantes:

- être un terme caractérisé par le locuteur comme n'étant pas endémique, par exemple *galeper* à 117 Lemoa, 123 Mañaria (qui le considère d'extraction guipuscoane), tout comme 124 Otxandio. 130 Arrazola le refuse aussi. Si le locuteur assigne à un terme proposé la qualité 'terme non endémique' le mot en question reste dans la base, sans accéder à la carte.

- ou bien être un terme connu par le locuteur dans son parler mais ayant un autre sens: c'est ci-dessus le cas de *poxpolin* quand il signifie pour le locuteur 'onomatopée du chant de la caille' laquelle caille n'est pas appelée *poxpolin*. Dans ce cas aussi, c'est la base de données qui conserve cette information et peut en fournir la transcription phonétique quand l'enquêteur l'a jugé intéressant. Fidèle à l'onomasiologie, la carte n'est pas accessible au terme qui n'a pas le sens correspond à l'intitulé de la carte.

C'est ainsi qu'à Gizaburuaga (point 133) le témoin refuse la proposition *poxpolin*. Urretxu (point 311) le refuse et le considère comme propre à Elgeta, commune qui n'est pas un point d'enquête de l'atlas mais sensée représenter le seuil du biscayen pour le locuteur guipuscoan à Urretxu. Azpeitia (point 316) et Aia (point 322) refusent aussi *poxpolin*.

A Deba (point 309), au terme proposé *poxpolin* le témoin déclare *Kantáu izán du, bañá esám bérari ez; "poxpólin kantátzē uená, galéperrá".... / Elle a un chant, mais on n'appelle pas ainsi l'oiseau lui-même, l'oiseau qui chante en faisant poxpolin c'est la perdrix (galeperra)*. Crucial est ce témoignage comme à Bergara mais plus clair. La forme du chant prêté à l'oiseau ne sert pas à le désigner. La carte ne rend pas compte du glissement. Ezkio-Itsaso (point 314) va dans le même sens que Deba et glose ainsi la proposition *poxpolin*: *Beónek ála kantátzen do..., len garí artén da... / Ah elle chante ainsi (poxpolin) autrefois on la trouvait dans les blés*. Ici non plus la carte ne peut le dire. Même phénomène à 318 Beizama *Kantátzeulakó : poxpóliñ! / On l'appelle ainsi (poxpolin) parce qu'elle chante poxpóliñ!* ou bien encore à 330 Andoain: *Poxpolin! poxpolin! kantatze ulako. Parce quelle chante poxpolin poxpolin*. 321 Errezil présente la même remarque atténuée par le mode interrogatif en commentant ainsi *poxpolin*: *áren (galeperraren) kantúa ez al dá poxpóliña?.. / Mais poxpolina n'est-ce pas le chant de la caille?* Plus loin vers l'ouest en Navarre on trouve la même observation à 403 Dorrao: *lóxpoló/ kantátzen dó ta... / elle chante poxpolo*, où d'ailleurs la forme *poxpolo* s'éloigne de *poxpolin*.

A Markin-Etxebarri (point 135) les témoins produisent : - *Garíxe eráitxe sanién da, eóte sán áskó, ta garí soluéen da ikúste sám pospóliñé; kantáu (egiten du) : txik txitxik! pospóliñ!!!.... - Garí botién botién ta botiem bós (esaten zuela/pospólin edo galeperrak/ -dio-), botiem bós, botiem bós, botiembós...., sárrak esándako gausie (da)... / Du temps où on semait du blé il y en avait alors beaucoup et dans les emblavures on voyait la caille (poxpolin) et elle chantait txik txik txik pospóliñ! Dans le blé botiem bos botiem bos botiem bos (chantait la caille), c'est une chose que*

*disaient nos anciens*. Ce texte se distingue des autres puisque outre notre habituel *poxpolin* on y trouve la séquence plus longue *txik txik pospolin* d'une part et celle nouvelle *botien bos* que nous ne savons pas interpréter. Faute de cette interprétation, nous ne pouvons en faire un mimologisme. Et Orexa (point 332) fournit aussi: *kra! kra! taatzea poxpoli poxpoli / Kra! Kra! Puis de nouveau poxpoli* avec une nouvelle.

En conclusion sur cet espace A3, nous pouvons donc voir que *poxpolin* qui n'est pas présent sur la carte, sauf sur les points (six en Guipuscoa) où il est accepté comme terme proposé, est cependant connu des locuteurs et permet de mieux comprendre que *poxpolin* soit désignation dans l'espace A2. Et c'est la base de données qui informe en complément de la carte. Nous insistons sur le fait que c'est surtout dans cette aire où *galeper* est réponse cartographiée que nous avons obtenu des informations sur *poxpolin* comme onomatopée du chant de l'oiseau.

Nous avons peu à dire sur la zone où la réponse est *ttorttola, txortxolo*, formée par quatre villages sur un espace continu en Navarre, six avec Zilbeti peu éloigné et avec Féas-Esquiule à l'extrême ouest. Le fait qu'à Suarbe (point 143) le témoin refuse *galeper* proposé en le glosant comme propre aux Guipuscoans plaide en faveur de la réponse obtenue à tenir pour endémique.

Nous consacrons maintenant quelques lignes à la zone massive orientale dans laquelle c'est l'emprunt *kalla* qui est seule réponse et le terme *galeper* n'y a nulle part été accepté. A cette apparente banalité correspondent cependant un certain nombre d'informations intéressantes sur la base.

Les témoins prononcent l'onomatopée prêtée à la caille et sur cet espace ce n'est jamais *poxpolin*: nous en relevons au moins six: *kok kok!*; *tinkaka!*; *titi pitti!*; *tha tha tha!*; *pal-pala!* et *trrrrrrrtütütütü* dans les points d'enquête de Saint-Etienne de Baigorri, Esterençuby, Pagolle, Domezain, Zugarramurdi, Sainte-Engrâce, qui livrent donc une très grande divergence sur la perception ou la restitution du chant prêtée à cet oiseau.

On aura remarqué la présence à nouveau de Pagolle comme fournissant une onomatopée en plus du seul mimologisme relevé quant à notre oiseau dans l'enquête: il s'agit de *petedet*, 'paie tes dettes' déjà évoqué. Le texte est le suivant: *Titi pitti egiten dizü; pétedet erraiten ziela, paie tes dettes, erraiten ziela ogian petik, zorrak phakatzen beitziren ogiz. Haurrian entzüten nizün. / Elle fait titi pitti!, on dit qu'elle chantait petedet, paie tes dettes, qu'elle le disait depuis sous les blés, parce que les*



*dettes alors se payaient en blé, je l'entendais étant enfant.* La justification est trouvée par le témoin dans le fait que les dettes étaient généralement payées en... blé.

D'ailleurs le blé est toujours invoqué comme habitat caractéristique de la caille, sauf exception à Esterençuby où la caille vivait sous les fougères en zone de montagne loin des emblavures et que les bergers entendaient à la saison de la tonte des ovins. Le texte dit: *Erran dautzut lehen e ibili dugu solasa amelarriain gainetik Errozatetik behiti badela halga, dena halgaden mendia, pheza handia bada, orenerdi bide bada eta lehen, oai ardiak moztuik joaiten dira, leheno ekhartzen gintien Errozatetik ardiak mozteat, jautsi gintien etxeat, bethi ardi mozteko sasoina uztailan zen, uztailain hamar irian, eta delako halga mendi hori trabesatzen gindielarik, zonbeit aldi ikhusten gintien, aitzenzintien ardien zangopetan kaila-umiak tinkaka ari zirenak, ono etziren airatzen hek eta hanitz han ziren belharren azpian. Berekantia badu horrek, "kalla kantuz ogipetik" aitzen zen. Ezta ogiik eta lehen iatze sailetan kalla munnahi bazen / Je vous l'ai dit tout à l'heure et nous en avons parlé au sujet de la pierre qui roule dans le précipice depuis le haut du pic d'Errozate, plus bas il y a une plante appelée 'halga', une montagne remplie de 'halga, il y a un vaste espace, d'une demi-heure de marche et autrefois, maintenant les brebis partent en estive après avoir été tondues, mais autrefois nous ramenions les brebis depuis Errozate pour les tondre, nous descendions à la maison, la saison pour les tondre était en juillet, autour du dix, et quand nous traversions cette fameuse montagne de 'halga', parfois nous les voyions, vous entendiez sous les pattes des brebis les petits des cailles qui étaient en train de chanter tinkaka!, eux ne n'envolaient pas encore, et il y en avait beaucoup sous l'herbe. Elle a son chant, la caille, nous chantions 'la caille chante depuis les blés'. Il n'y a plus de blé mais autrefois dans les fougères il y en avait partout.*

Dans les autres points d'enquête, les témoins établissent aussi un lien entre la disparition de la production de blé et la raréfaction de la caille. Domezain indique par exemple: *Lehen bazen biziki ogi gehio oai beno, guk iten duun oono ogi,hiu ektara et'erdi, eta humiak iten dauzküte hor ogin barnen, thathathatha ai da ederki / Autrefois il y avait beaucoup plus de blé que maintenant, nous on fait encore du blé, trois hectares et demi et les cailles nous font des petits dans les blés, ils font thathathatha, c est joli.*

Le lecteur bascopphone aura remarqué que quand nous parlons du blé, le terme qui est utilisé en domaine oriental est *ogi* et non pas *gari*. En domaine oriental en

effet (zone de *kalla*), le même mot *ogi* sert à désigner aussi bien le ‘pain’ ordinaire que le ‘blé’. Le terme *gari* n’y est pas connu, et ceci est à mettre en relation avec l’absence du composé *galeper* sur cette aire. Nous ne ferons pas ici l’excursion sur la dépendance ou indépendance des aires lexicales mais disons de manière très simplifiée que l’aire de *kalla* ne connaît que le mot *ogi* pour blé.

Enfin, même si notre atlas ne se veut pas un outil de collecte de chansons et d’éléments phraséologiques, les témoins fournissent souvent des éléments ethnolinguistiques. Ici par exemple plus d’un témoin a donné les paroles d’une chanson populaire dont le premier mot est *kalla* la caille: parfois ils chantent les premières mesures, par exemple aux Aldudes, point 601. Le témoin de Macaye a donné la première strophe:

<i>Kalla kantuz / ogi petik</i>	<i>Sous les blés chante la caille</i>
<i>uztaila agorriletan/</i>	<i>En juillet et août</i>
<i>maitiaganik etxerakuan</i>	<i>Quand je regagnais mon logis revenant de chez</i>
<i>entzun izan dut frangotan</i>	<i>mon aimée</i>
	<i>Je l’ai eu maintes fois entendue.</i>

Répetons-le et même si ce n’est pas la fonction première de l’atlas de tels passages intéresseront le collecteur, non pas souvent comme texte inédit (le nombre des versions données montre que la chanson est très connue, à Arnéguy, Barcus, Irissary...) mais comme variation possible. La chanson ci-dessus se trouve dans les recueils de chansons tels que le *Cancionero Vasco* d’Azkue (Azkue 1921-1925, I-108). La date permet de constater que *kalla* n’est pas un emprunt des jeunes générations puisque présent déjà au XIX<sup>ème</sup> siècle.

La question ‘caille’ n’a pas donné lieu à beaucoup d’éléments phraséologiques dans notre enquête. Nous relèverons seulement une comparaison à Uhart-Cize (point 607): *Kalla bezala lanian artzia, xehatu artio esklabotasuna, azken piruraino / Se tuer au travail comme une caille, l’esclavage jusqu’à se détruire,, s’user jusqu’au dernier fil.*

Nous devons à la vérité de dire que ces éléments phraséologiques sont éparpillés dans toute la base sans être étiquetés en vue d’une recherche spécifique: ce travail devra être réalisé ultérieurement.

Les remarques métalinguistiques sur l’emprunt *kalla* n’ont pas été très nombreuses mais toutefois à Chéraute (point 708) le témoin indique: *Biarnesez e kalla*

dük / *Tu sais, en béarnais aussi on dit kalla*, sans que l'origine du terme ne soit stigmatisée explicitement.

Pour résumer les informations fournies dans l'aire de *kalla*, emprunt manifeste, que la carte pourrait assimiler à une morne plaine, nous avons donc pu mettre en lumière d'autres informations: une grande variété des onomatopées (toutes distinctes de *poxpolin*), un mimologisme en langue française, le recours à une chanson traditionnelle qui signale que l'emprunt n'est pas tout récent ou dû à la diglossie récente, le lien entre la caille et le blé qui invite le lecteur à consulter la carte 'blé', quelques éléments de phraséologie et des remarques métalinguistiques.

### 3. Conclusions

Il est temps de conclure. Nous pensons après ce rapide survol avoir montré que l'atlas est une mine de renseignements dont notre cartographie veut rendre compte de la manière la plus visible possible (format réduit, simplification du fond de carte, lissage des réponses par lemme et recours à la couleur, détail des formes phonétiques sur un responsable). Mais pour aussi détaillée que soit la carte, la masse d'autres données recueillies et leur degré de nuance échappent à la cartographie et peuvent intéresser le lecteur attentif à la variation et commentaire des témoins dans les domaines de la lexicologie, de l'ethnolinguistique, de la conscience linguistique. En dernière analyse, l'auditeur scrupuleux aura toujours le recours du phonogramme comme source des données exploitées. La matière protéiforme qui apparaît ainsi nous rapproche des conditions de l'enquête et du monde de nos témoins.

Nous pensons produire ainsi un apport positif et la publication prochaine de ces travaux nous dira si le public trouve intérêt à la méthodologie suivie. Dans le respect de l'objet linguistique singulier, nous essayons de gagner en clarté, commodité, lisibilité... pour que les atlas linguistiques soient davantage utilisés en lexicologie en particulier et en linguistique en général.

Nous avons parlé des informations conservées sur la base de données. La cartographie utilisée permet de mieux mettre en lumière les caractéristiques plus importantes qui se cachent dans la forêt des réponses : faits intrabasques (les variantes d'un même lexème), présence d'un pluricentrisme (*zeken-zokor*... 'taureau'), voie d'introduction des emprunts. Elle devrait aussi permettre de mieux mettre en valeur l'insertion particulière de la langue basque dans son espace pyrénéen au contact des

langues romanes : emprunts, calques, motivations similaires, ce qui permet de faire de l'atlas basque un outil qui peut être utile à la linguistique générale, en particulier à la sémantique, à la cognition, mais aussi à l'histoire culturelle de l'Europe occidentale.

### Bibliographie:

- Aurrekoetxea, G. / X. Videgain, 1989, "L'atlas linguistique du Pays Basque. Euskal Herriko Hizkuntza Atlas", *Bulletin du Musée Basque*, 1. trim., 495-510
- Aurrekoetxea, G. / X. Videgain, 1993, "Euskal Herriko Hizkuntza Atlas. Galdesorta / cuestionario / questionnaire", *Euskera* XXXVIII-3, 529-647
- Aurrekoetxea, G. / X. Videgain, 1994, "Historia y futuro del Atlas Lingüístico Vasco (EHHA)", in P. García Mouton (arg.), *Geolingüística. Trabajos europeos*, Biblioteca de Filología Hispánica, CSIC, Madril, 79-96
- Aurrekoetxea, G., 1986, "Euskal Herriko Hizkuntza Atlas (EHHA): inkesta metodologia eta ezezeko datuak", *Euskera* XXXI, 413-424
- Aurrekoetxea, G., 1992, « Euskal Herriko Hizkuntza Atlasaren informatizapena », in Aurrekoetxea, G. & X. Videgain (éd.), *Actes du Congrès International de Dialectologie*, IKER-7, Euskaltzaindia, Bilbao, 319-338.
- Azkue, R.M., 1921-1925, *Cancionero vasco*, Bilbao.
- Chambon, J.P., 1989: "Démimologiques: délocutivité et zoonymie dans le domaine galloroman", *Bulletin de la Société de Linguistique*, 1, 81-110.
- Dinguirard, J.C., 1976 : *Ethnolinguistique de la haute vallée du Ger*, Thèse Université Toulouse-le Mirail (dir. : J. Allières).
- Palay, S., 1991 [1974]: *Dictionnaire du béarnais et du gascon modernes (Bassin aquitain)*.
- Ravier, X., 1965: "Les données négatives dans l'ALG", *RliR*, vol. 115-116, 262-274.
- Tanguy, B., 1984 "Analyse lexicographique et cartographique de l'Atlas Linguistique de la Basse-Bretagne de Pierre Le Roux", *Studi*, n° 25
- Van Gennep, A., 1934: "Contribution à la méthodologie du folklore", *Lares*, V, 1, 20-34.
- Videgain, X., 1992, « Lexikoaren bilketaren baldintza zenbait », in Aurrekoetxea, G. & X. Videgain (éd.), *Actes du Congrès International de Dialectologie*, IKER-7, Euskaltzaindia, Bilbao, 559-576.